Am J Hum Genet 35:1161–1178, 1983

The Transmission of Schizophrenia under a Multifactorial Threshold Model

M. McGue,[1] I. I. Gottesman,[2] and D. C. Rao[1,2]

SUMMARY

Family studies of schizophrenia have reported elevated rates of both definite and definite-plus-probable schizophrenia among the relatives of definite schizophrenics. These elevated rates imply a strong association between the two forms of diagnosis and suggest some form of familial transmission. Here we have used recently developed maximum likelihood methods to investigate this association and characterize the nature of the familial transmission.

Results indicated that although the two forms of diagnosis were strongly related, they could not be considered alternative manifestations of a single liability distribution. Heritability estimates for either form of diagnosis were comparable ($h^2 = .668 \pm .052$ and $c^2 = .191 \pm .038$ for definite while $h^2 = .628 \pm .073$ and $c^2 = .236 \pm .106$ for definite-plus-probable), although cultural transmission (i.e., c^2) was statistically significant only for definite-plus-probable. For either form of diagnosis, residual twin resemblance was statistically significant and could not be explained in terms of the effects of genetic dominance. These results are comparable to those of an earlier analysis based upon a similar data set. Finally, the statistical correction used to adjust for between-study heterogeneity in morbidity risk figures did not noticeably alter the parameter estimates.

INTRODUCTION

Schizophrenia is a common familial psychiatric disorder with a lifetime risk of approximately 1% by age 55 [1, 2], but the exact nature of its transmission

Received November 24, 1982; revised February 18, 1983.

This work was supported in part by grant GM-28719 from the National Institutes of Health and grant MH-31302 from the National Institutes of Mental Health.

[1] Division of Biostatistics, Washington University School of Medicine, Box 8067, 4566 Scott Avenue, St. Louis, MO 63110.

[2] Departments of Psychiatry and Genetics, Washington University School of Medicine, Box 8134, 4566 Scott Avenue, St. Louis, MO 63110.

remains uncertain [3–8]. The progression of hypothesized models for the mode of transmission of schizophrenia has closely mirrored advances in methodological and computational techniques. Early researchers hypothesized that schizophrenia was a completely penetrant recessive disorder [9], but such a simple mode of transmission could be rejected by the finding of substantially less than 100% affected offspring from the matings of two schizophrenics together with equal risks to siblings and offspring of schizophrenics [10–12].

Subsequently, more general single-locus models have been hypothesized for the transmission of schizophrenia [13]. Under the generalized single major-locus model, the transmission of schizophrenia is due to the segregation of two alleles at a single locus. Attempts to fit these models to family or pedigree data have led to either equivocal [6, 13] or disconfirmatory [14] results. In a recent comprehensive assessment of the evidence for and against single-locus transmission, O'Rourke et al. [7] concluded that such generalized single-locus models were not consistent with findings from family, twin, and dual-mating studies of schizophrenia; as was the case with the simple completely penetrant recessive hypothesis that preceded it, the weight of empirical evidence did not support the generalized single-locus model of schizophrenia. This is not to conclude that there is no major gene (or genes of large effect) contributing toward the risk for developing schizophrenia, but only that such a major gene, should it exist, cannot be the only factor in the transmission of schizophrenia. Other factors that could be contributing to the transmission of schizophrenia, but are not taken into account under the generalized single-locus models, would include polygenic and familial environmental effects.

The finding that a single gene is not sufficient to explain the transmission of schizophrenia implicates additional or alternative sources of transmission. Alternatives to the generalized single-locus models begin by assuming schizophrenia to be a threshold character [15, 16]; that is, underlying the categorical phenotype (i.e., diagnosis) of schizophrenia is an assumed continuous liability toward developing schizophrenia. Everyone possesses some liability for developing schizophrenia, but the expression of schizophrenia occurs only when an individual's combined liability exceeds some fixed threshold value along the liability continuum. Specific models for the transmission of schizophrenia as a threshold character can be characterized by those factors assumed to contribute toward the liability for developing schizophrenia.

The mixed model of Morton and MacLean [17] allows for major gene, polygenic, and environmental effects upon liability and thus includes sources of transmission other than a single locus. Although the mixed model is the most general model of transmission available and has proven useful in the study of many quantitative as well as qualitative characters, many of the complexities of the schizophrenic phenotype (e.g., variable age of onset and low proportion of multiplex families) make it difficult to analyze using current programs for mixed model analysis. Furthermore, the power of a mixed model analysis to identify major gene effects using qualitative data alone is reduced relative to an analysis based on quantitative data. This, perhaps, explains the equivocal findings of Carter and Chung [18],

who failed to resolve major gene vs. polygenic effects statistically in their mixed model analysis of 365 schizophrenic pedigrees.

Polygenic transmission represents an alternative to single-gene models. Falconer [15], in his study of congenital malformations, common genetic diseases, and other threshold characters, considered only polygenic transmission. Gottesman and Shields [16] were the first to apply this polygenic model to schizophrenia.

Multifactorial threshold models represent major alternatives to the single-locus models. Under a multifactorial model, familial transmission is due to multiple genetic (i.e., polygenic) and multiple environmental (familial) factors. Although no attempt is made to characterize major gene effects under the multifactorial models, their existence does not necessarily invalidate applications of the model, an issue that we will address later.

Attempts to study the transmission of schizophrenia under the multifactorial threshold model have, until recently, been hindered by the lack of efficient and powerful methods of statistical analysis, difficulties in defining a data set that would allow resolution of the various transmitted factors, and finally the effects of varying diagnostic standards on the familial morbidity risk figures (see, for example, [19, 20]). Recently, Rao et al. [8] and Rice et al. [21] adopted a general likelihood approach to develop methods for the analysis of the transmission of threshold characters under multifactorial models that in the former case included separate components for both polygenic and familial environmental transmission, and in the latter case, included a single component for both types of transmission. Rao et al. applied these techniques to an analysis of European family studies of schizophrenia and concluded that the transmission of schizophrenia could be accounted for under a liability threshold model with estimates of the genetic and cultural heritabilities for liability given by .707 ± .077 and .203 ± .120, respectively.

Here we have updated the Rao et al. European family data set and addressed some specific issues that remain unresolved in the earlier analysis. First, can we increase our ability to characterize the transmission of schizophrenia by pooling information across several studies, or does pooling add noise to the analysis? Second, as there are varying diagnostic standards depending upon the severity of the disorder, can information about individuals affected with less severe forms of schizophrenia be used in our analysis, and, if so, can our model of transmission account for this gradient of severity? In the first section, we describe available European family studies of schizophrenia and define a data set to be used in subsequent analyses. In the second section, we introduce the model and associated method of analysis. The results are presented in the third section and discussed in the final section.

FAMILY STUDIES OF SCHIZOPHRENIA

The genetic and cultural transmission of schizophrenia can be resolved through an analysis of family study results [8]. In table 1, we report the pooled results of relevant twin and family studies of schizophrenia, all of which have been undertaken in Western Europe. European investigators, working within a relatively

TABLE 1

CHARACTERISTICS OF THE EUROPEAN FAMILY STUDIES OF SCHIZOPHRENIA AS A FUNCTION OF THE DIAGNOSIS IN RELATIVES OF DEFINITE SCHIZOPHRENICS

	DEFINITE DIAGNOSIS					DEFINITE-PLUS-PROBABLE DIAGNOSIS				
	BZ	Affected (%)	r*	No. studies	Heterogeneity† χ^2	BZ	Affected (%)	r*	No. studies	Heterogeneity† χ^2
Spouses	399	1.0	.026	4	4.60	399	2.26	.128	4	5.18
Children	1678.6	9.35	.439	7	17.27‡	1254.3	12.84	.499	5	6.19
Siblings	7523.2	7.30	.381	10	44.32‡	7628	9.26	.418	10	52.62‡
MZ twins	106	44.3	.853	3	1.05	261	45.6	.854	5	3.88
DZ twins	149	12.08	.501	3	2.33	329	13.7	.515	5	8.87
Half-siblings (reared by common parent)	442.5	2.94	.198	3	7.09‡	267	5.99	.317	2	0.54
Nieces-nephews	3965.5	2.65	.179	6	15.65‡	2973	3.46	.202	3	8.43‡
First cousins	1600.5	1.56	.092	3	2.30	1600.5	2.44	.141	3	0.38
Grandchildren	739.5	2.84	.192	5	3.64	382	4.97	.278	3	2.11

* Tetrachoric correlations calculated using lifetime risk of 0.85% for definite and 0.99% for definite-plus-probable.
† Degrees of freedom for heterogeneity χ^2 is 1 less than the no. studies.
‡ Heterogeneity χ^2 significant at $P < .05$.

nonmobile, homogeneous society with a strong tradition of genetic epidemiology, have enjoyed better access to and cooperation from the relatives of schizophrenics than their American counterparts [2]. Consequently, we will restrict our analyses to Western European studies alone. The compilation given in table 1 represents all published Western European family and twin studies of schizophrenia with the following exceptions. First, adoption studies [22] have employed unique selection and diagnostic criteria that preclude meaningful pooling with the other family studies. Second, as parenthood involves considerable selection for mental health [23], the risks to parents would not be comparable to other risks and thus are not included. Third, as the concept and diagnosis of schizophrenia is a relatively new one [24], risks to grandparents and other ascendants of schizophrenics are likely to be unreliable, subject to underascertainment and are thus not included. Finally, as described by Gottesman and Shields [2], due to unique sampling characteristics, a few individual family studies were deleted as being noncomparable to other family studies. Usually such studies used probands who were affected with other psychiatric disorders (e.g., alcoholism, mental retardation) in addition to their "schizophrenia." Application of these criteria resulted in a different data set from that used by O'Rourke et al. [7] in their analysis of the generalized single-locus model.

Table 1 reports the total age-corrected sample sizes (BZ), the corresponding risk of affection in relatives, the tetrachoric correlations (calculated using lifetime risks for both forms of diagnosis given by Slater and Cowie [10]), the number of independent studies pooled, and the χ^2 statistic used to test the significance of the between-study heterogeneity in the rates of affection. In all cases, probands were definite schizophrenics. The table provides this information for both "narrow" or "definite" diagnoses in relatives of probands and for "wide" or "definite-plus-probable" diagnoses. Except for twins, all risk figures have been age-corrected. The substantial correlation between age of onset for twins both implies little need for adjustment, given the follow-up intervals reported for cotwins, and invalidates standard age-adjustment methods (cf. [25–27]). Consequently, twin rates remain unadjusted. All concordance rates for twins were probandwise concordance rates.

For definite diagnosis, we find statistically significant between-study variability in four of the nine classes of relatives. It should be emphasized that, despite the significance of the χ^2 statistics, no single investigation has a substantial effect on the pooled risks reported in table 1. For example, deleting any single report on the siblings of schizophrenics would change the present risk of 7.3% for definite schizophrenia by no more than 0.35%. Although there are fewer significant heterogeneity χ^2's when we move to the definite-plus-probable risks (only two of nine), it would not be appropriate to conclude that allowance for a probable diagnosis reduces between-study variability. If we take the sum of the χ^2 statistics divided by the sum of their degrees of freedom (χ^2/df) as a measure of heterogeneity (a measure that is more sensitive than a simple count of the number of significant χ^2's), then we find that both diagnoses result in comparable levels of heterogeneity. For definite diagnoses, the total, over the nine classes of relatives, χ^2/df is 2.81, while for definite-plus-probable, the χ^2/df is 2.85. In a similar manner, it can be

shown that rates reported for remote classes of relatives (i.e., second and third degree) are not more heterogeneous than rates reported for first degree relatives. The latter finding is at odds with the notion that ascertainment of remote relatives is less uniform across studies than ascertainment of first degree relatives, at least for those studies of descendants and collaterals of schizophrenics.

Given the existence of between-study variability, there are several options one could take in defining a data set for analysis. First, one could take the position that the existence of any heterogeneity precludes meaningful pooling of the data, and, consequently, that all analyses should be restricted to single investigations. Alternatively, we could delete all data on those classes of relatives for which there was significant between-study heterogeneity with the hope that the resulting data set would be both homogeneous and allow for the resolution of the transmission of schizophrenia. A final alternative, less wasteful of information, is to correct statistically for the effects of heterogeneity upon the variances of the pooled statistics by using some appropriately defined scaling factor.

Although we are able to take any of these approaches with the present data set, we prefer the third strategy as it makes maximal use of the information available. It can be shown (APPENDIX) that the primary statistical effect of between-study heterogeneity is to inflate the variances of the pooled statistics, this inflation being roughly proportional to the heterogeneity χ^2 divided by its degrees of freedom. For example, the combined risk derived by pooling the individual risks observed in k homogeneous studies on a total sample of size N would have a variance proportional to N^{-1}, while if the k studies were heterogeneous, the variance of the pooled risk would be approximately proportional to $(N(k-1) \div \chi^2)^{-1}$. Consequently, the additional source of variance in the summary risks due to between-study heterogeneity can be accounted for by scaling the total sample sizes for heterogeneous classes to be $N \div (\chi^2/\mathrm{df})$ rather than N. Rao and Morton [28] and Rao et al. [29] used a similar argument for scaling sample sizes in their analysis of the familial resemblance for IQ. In the analyses that follow, unless otherwise indicated, we have adjusted the sample sizes only for those classes where we found significant between-study heterogeneity. Note that this adjustment does not affect the rate of affection or the tetrachoric correlation. We will compare results using this correction to results based upon some form of deletion of heterogeneous data to assess the effects of this adjustment.

In most European investigations, two forms of diagnosis have been used for relatives of schizophrenics; all probands had a definite diagnosis. A definite diagnosis was made whenever a relative met the standard (i.e., "textbook") diagnostic criteria for schizophrenia. A probable diagnosis was made whenever an individual did not meet standard diagnostic criteria, but, nonetheless, had most of the signs and symptoms of schizophrenia, was psychotic, and was more likely to be schizophrenic than to have an affective psychosis [11, 19, 27]. Given the notion of an underlying liability to schizophrenia, we might hypothesize that the two forms of diagnosis represent different thresholds along the same liability distribution (fig. 1; [30]). We can test whether a single liability distribution can account for the two forms of diagnosis, as, if so, we would expect the tetrachoric correlations calculated using either form of diagnosis to be equal. In table 2, we

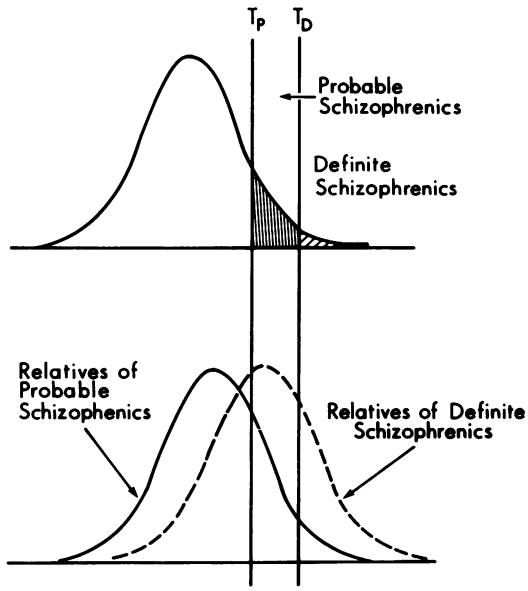

FIG. 1.—Transmission of definite and definite-plus-probable schizophrenia under a two-threshold model. Definite schizophrenia occurs whenever an individual's combined liability exceeds T_D; definite-plus-probable schizophrenia occurs whenever an individual's combined liability exceeds T_P. Note that threshold values are not to scale with the lifetime risks of 0.14% and 0.85% used in the analysis.

give the tetrachoric correlations for each form of diagnosis as well as the estimate of the common (polychoric) correlation and the χ^2 statistic used to test the equality of the two correlations. Correlations were estimated by the method of maximum likelihood. For the proportion of relatives with definite schizophrenia, the likelihood was taken to be the integral of a singly truncated normal distribution, and for the proportion of relatives with a probable diagnosis, the likelihood was taken to be the integral of a doubly truncated normal distribution. In calculating correlations, the lifetime risk for definite schizophrenia was taken to be 0.85%, while the lifetime risk for definite-plus-probable was taken to be 0.99%. There is substantial support for 0.85% being the risk for definite schizophrenia, including Slater and Cowie's compilation of 17 population-based surveys undertaken in Western Europe [10]. The lifetime risk for a definite-plus-probable diagnosis would appear to be in greater doubt. Only three of the above-mentioned population surveys reported rates for definite-plus-probable, the average being 0.99% $\pm$ 0.098 [31–33]. Correlations were computed using only those studies in which both forms of diagnosis were used; consequently, the samples reported in table 2 represent a subset of those given in table 1. The χ^2 statistic in table 2 is the likelihood ratio test statistic testing the hypothesis of single common correlation for both forms of diagnosis against the general hypothesis of two separate correlations. As can be seen, the two-threshold model is rejected; the overall χ^2 for the fit of the two-threshold model is 49.93, which on 9 df is highly significant ($P \doteq 0$). Comparing the correlations for the intermediate category against those for the extreme category, we note that the former are uniformly higher, suggesting

TABLE 2

TETRACHORIC AND POLYCHORIC CORRELATIONS AS A FUNCTION OF DIAGNOSIS IN THE RELATIVES OF DEFINITE SCHIZOPHRENICS

Relationship to proband	Definite only	Probable only	Common for definite-plus-probable	χ^2 (1 df) test for homogeneity of correlations
Spouses	.026	.479	.136	9.23*
Children	.440	.778	.598	1.03
Siblings	.367	.731	.509	4.76*
MZ twins	.853	.995	.972	5.31*
DZ twins	.501	.917	.703	5.16*
Half-siblings	.279	.533	.375	2.21
Nieces-nephews	.199	.201	.198	3.44
First cousins	.092	.378	.158	13.57*
Grandchildren	.225	.548	.323	5.22*
Total (9 df)				49.93*

NOTE: Tetrachoric correlations for definite only calculated using lifetime risk of 0.85%; tetrachoric correlations for probable only calculated using lifetime risk of 0.14%; polychoric correlations calculated assuming two grades with lifetime risk of extreme grade given by 0.85% and intermediate grade given by 0.14%.

* Test of equality of two tetrachoric correlations significant at $P < .05$.

that a higher population risk figure for the intermediate category might result in a fit of the two-threshold model. As there already existed some question regarding the accuracy of the 0.99% figure, we tested the two-threshold model for three additional rates for definite-plus-probable: 1.17%, a rate suggested by Slater and Cowie [10]; 1.25%, an intermediate rate; and 1.39%, the highest rate reported for definite-plus-probable from the three population surveys [31]. In all cases, the two-threshold model could be rejected at $P \leq .05$, implying that the familial data rather than specification of the population risk figures led to rejection of the two-threshold model.

As the risks for probable diagnoses among relatives of definite schizophrenics is substantially greater than the population risk of 0.14% (i.e., 0.99%–0.85%), we are able to conclude that there is a strong, albeit not perfect, association between the liabilities underlying the two forms of diagnosis. Smith [34] reported similar results with respect to the relationship between early- and late-onset diabetes and the relationship between anencephaly and spina bifida, and is able to quantify the degree of overlap between the liabilities for the disorders because he had risk figures on relatives of both types of probands. Unfortunately, we do not have information on the relatives of individuals with a probable diagnosis of schizophrenia, and, consequently, we are unable to quantify the strength of the association between the two liabilities [30]. As a result, in all subsequent analyses, the two forms of diagnoses (i.e., definite and definite-plus-probable) will be treated separately.

MULTIFACTORIAL THRESHOLD MODEL

Under the present formulation, the transmission of schizophrenia is a result of the transmission of the liability to schizophrenia. In figure 2, we give a schematic representation of the path model that we propose for the transmission of liability within a nuclear family. Liability is assumed to be an additive function of the

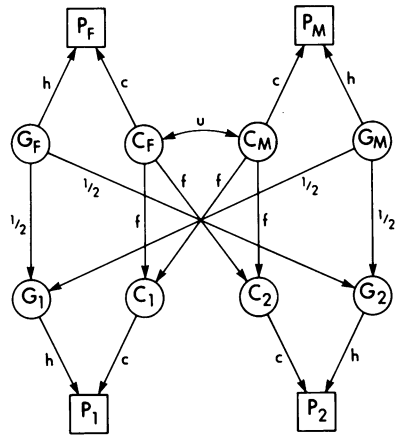

FIG. 2.—Path diagram for the transmission of schizophrenia in nuclear families. *G*, *C*, and *P* denote genotype, transmissible environment, and phenotype, respectively. *Subscripts F*, *M*, *1*, and *2* denote father, mother, and two children, respectively.

effects of an individual's genotype (G), transmissible family environment (C), and an uncorrelated residual environment, which, for convenience, is not pictured. Only additive genetic effects are considered (dominance and epistasis are assumed to be zero; cf. [35]). Transmissible family environmental effects reflect those environmental factors that are potentially transmissible from one generation to the next (e.g., social class, diet, etc.). Residual environmental effects are factors that family members do not share in common (e.g., head trauma, birth complications, etc.). Although the interaction of genotype and familial environment is assumed to be zero, the interaction of either factor with residual environmental factors would inflate the variance attributable to residual effects but not otherwise affect the analyses. Within nuclear families, there are four parameters in the model; the genetic heritability, h^2; the cultural heritability, c^2; the effect of parental environment upon offspring's environment, f; and the correlation between the transmissible environments of spouses, u. In addition, we have added a fifth parameter specific to the analysis of twin data. This parameter, t^2, is that part of the total residual variance that is common to either MZ or DZ twins and reflects all factors such as greater similarity in the trait-relevant environments (prenatal and/or postnatal) and the existence of nonadditive genetic effects that would lead to greater resemblance of twins as compared to ordinary siblings. Finally, it should be emphasized that even though there is no explicit allowance for major gene effects in the multifactorial threshold model, they are not necessarily excluded. If a major gene is affecting liability, then h^2 would include additive variation at the major locus while dominance at the major locus would probably lead to an elevation of both t^2 and f.

The expected correlations between liability values of any two members of a nuclear family, twins, or more distant relatives can be expressed in terms of the five parameters of the model as shown by Rao et al. [36]. For the classes of relatives studied here, the expressions for the expected correlations are given in table 3.

TABLE 3

COMPARISON OF OBSERVATIONS AND EXPECTATIONS UNDER THE GENERAL FIVE-PARAMETER MODEL

RELATIONSHIP	EXPECTED CORRELATION	PERCENTAGE DEFINITE		PERCENTAGE DEFINITE AND PROBABLE	
		Observed	Expected	Observed	Expected
Spouses	c^2u	1.00	1.19	2.26	2.57
Children	$\frac{1}{2}h^2 + c^2f(1 + u)$	9.35	7.85	12.8	12.1
Siblings	$\frac{1}{2}h^2 + 2c^2f^2(1 + u)$	7.30	6.88	9.26	10.1
MZ twins	$h^2 + 2c^2f^2(1 + u) + t^2$	44.3	43.3	45.6	45.1
DZ twins	$\frac{1}{2}h^2 + 2c^2f^2(1 + u) + t^2$	12.1	12.4	13.7	14.3
Half-siblings	$\frac{1}{4}h^2 + 2c^2f^2(1 + u)$	2.94	3.00	5.99	4.99
Nieces-nephews	$\frac{1}{4}h^2 + 2c^2f^3(1 + u)^2$	2.65	2.67	3.46	3.84
First cousins	$h^2/8 + 2c^2f^4(1 + u)^3$	1.56	1.55	2.44	2.10
Grandchildren	$\frac{1}{4}h^2 + c^2f^2(1 + u)^2$	2.84	2.81	4.97	4.34

Likelihood Formulation and Methods of Analysis

The details of the likelihood formulation are given by Rao et al. [8]. For each of the nine types of relationships, assume we observed A_i affected relatives and U_i unaffected relatives ($i = 1, \ldots, 9$). For heterogeneous classes, both A_i and U_i have been scaled as described above. The observations for one relationship are assumed to follow a binomial distribution with the total log likelihood for the nine types given by

$$\ln L = \sum_{i=1}^{9} [A_i \ln q_i + U_i \ln (1 - q_i)] ,$$

where q_i is the probability that a relative of type i is affected. Assuming that the liability distribution for any two relatives is bivariate normal, then q_i will be a function of the correlation between relatives of type i, and, consequently, ln L can be written as a function of the expected correlations (i.e., as functions of the parameters of the model). The formulation of the likelihood function given above is strictly valid only when we observe relatives of one class per proband. Observations on larger sets of relatives would introduce nonbinomial error, although no bias.

In this way, we can obtain the maximum likelihood estimates of the parameters by maximizing ln L and use the residual ln L value to test hypotheses about the nature of transmission. Let $\ln L_1$ be the residual value when $\alpha + \beta$ parameters are estimated and $\ln L_2$ be the residual value when only α of the parameters have been estimated, the other β parameters being fixed under a null hypothesis. Then the likelihood ratio test of the null hypothesis on the β parameters is given by $\chi^2 = 2(\ln L_1 - \ln L_2)$, which asymptotically follows a χ^2 distribution on β df.

Analysis consisted of first fitting the general five-parameter model and then several submodels that allowed tests of the hypotheses of no genetic heritability ($h^2 = 0$), no cultural heritability ($c^2 = f = u = 0$), no special twin resemblance ($t^2 = 0$), and no marital resemblance ($u = 0$). In all analyses, the rate of definite

schizophrenia was fixed at 0.85%, while the rate of definite-plus-probable was fixed at 0.99%.

All analyses were performed using ATRIBUTE, a FORTRAN program developed on the Harris computer [8, 35]. The mixed homogamy model of Rao et al. [36], with minor changes [29], has been implemented in ATRIBUTE, which accepts data on 24 types of relatives.

RESULTS

In table 4, we give the results of the hypothesis tests and the parameter estimates for both forms of diagnosis. For definite diagnosis, there is a very good fit of the general model to the data ($\chi^2_4 = 2.43, P = .66$). The estimated genetic heritability is large ($h^2 = .668 \pm .052$), while the estimate of cultural heritability ($c^2 = .191 \pm .038$) and special twin resemblance ($t^2 = .140 \pm .034$) are both moderate. A model that assumes no genetic transmission does not fit the data and leads to a substantial increase in the residual χ^2 statistic ($\chi^2_1 = 42.56, P \doteq 0$). In contrast, cultural transmission is not statistically significant when tested using the likelihood ratio test ($\chi^2_3 = 3.12, P = .37$). We find evidence in favor of additional resemblance unique to twins ($\chi^2_1 = 11.22, P = .001$). Finally, we do not find evidence for marital resemblance ($\chi^2_1 = 0.51, P = .48$).

The results based upon the definite-plus-probable diagnosis, although similar, are not in complete agreement with those found with the definite diagnosis. Again, the general model fits the data well ($\chi^2_4 = 4.19, P = .38$). The estimate of genetic heritability, although slightly less than before, is in close agreement with that found with the definite diagnosis ($h^2 = .628 \pm .073$). Both cultural heritability ($c^2 = .286 \pm .106$) and assortative mating ($u = .516 \pm .257$) are estimated to be more pronounced, while special twin resemblance is estimated to be lower ($t^2 = 0.086 \pm .025$). A different picture emerges from the results of the hypothesis tests for definite-plus-probable diagnosis. In this case, all effects are significant. Not allowing any genetic transmission leads to a significant increase in the residual χ^2 ($\chi^2_1 = 86.84, P \doteq 0$) as does no allowance for cultural transmission ($\chi^2_4 = 23.78, P \doteq 0$). Special twin resemblance as well as marital resemblance are significant ($\chi^2_1 = 4.68, P = .031$; and $\chi^2_1 = 9.51, P = .002$, respectively). Because there was some question regarding the validity of the 0.99% risk rate for definite-plus-probable, we also fit the general path model assuming a risk of 1.17%, the rate which, of the four tested, gave the best fit for the two-threshold model. The results were consistent with those based upon 0.99%, the general model fitting the data well ($\chi^2_4 = 3.58, P = .47$), with estimates of genetic and cultural heritability being given by $h^2 = .637 \pm .075$ and $c^2 = .279 \pm .106$.

To further investigate the fit of the general model and identify any major discrepancies, we compared the observed rates of affection with those expected under the general model in table 4. For definite diagnosis, the observed and expected rates are in close agreement. Only for children is there a discrepancy of as much as 1.0% between observation and expectation. Similarly, for definite-plus-probable, we find a good agreement between observations and expectations, although here the only relationship that demonstrated even a minor deviation from expectation was half-siblings. Additionally, in table 4, we find no systematic

TABLE 4

PARAMETER ESTIMATES AND TESTS OF HYPOTHESES UNDER THE GENERAL MODEL OF FIGURE 2

MODEL	DEFINITE SCHIZOPHRENIA							DEFINITE-PLUS-PROBABLE SCHIZOPHRENIA						
	χ^2	df	h^2	c^2	t^2	f	u	χ^2	df	h^2	c^2	t^2	f	u
General*	2.43	4	.668 ± .052	.191 ± .038	.141 ± .034	.256 ± .097	.246 ± .328	4.19	4	.628 ± .073	.286 ± .106	.086 ± .025	.364 ± .053	.516 ± .257
No genetic heritability ($h = 0$)*	44.99	5	0	.679	.321	.486	.107	91.03	5	0	.738	.262	.473	.263
No cultural heritability ($c = f = u = 0$)	5.55	7	.740	0	.112	0	0	27.97	8	.859	0	.0001	0	0
No environment unique to twins ($t = 0$)*	13.65	5	.785	.215	0	.089	.124	8.87	5	.763	.165	0	.326	.861
No marital resemblance ($u = 0$)*	2.94	5	.679	.187	.134	.272	0	13.70	5	.674	.278	.048	.440	0

* All analyses involve the constraint that $h^2 + c^2 + t^2 \leq 1.0$. Equality was attained in this case.

tendency for rates in remote relatives to be underpredicted, a finding at odds with a recent suggestion that schizophrenia is overdiagnosed in first degree relatives while underdiagnosed for more remote relatives [37].

DISCUSSION

In the case of a definite diagnosis, we found that the transmission of liability to schizophrenia could be accounted for by genetic factors only, although some source of special twin resemblance is necessary. Cultural transmission, should it occur, does not appear to be a major contributor to liability. This finding of marginal cultural transmission is consistent with attempts to identify specific environmental factors in schizophrenia [2, 38, 39]. At present, these investigations have yielded little in the way of specific contributors and have led to the conclusion that environmental factors in schizophrenia are idiosyncratic and random [2, 40]. Under the present formulation, these idiosyncratic effects are important and could contribute as much as 26% (1.0 − .74, see table 3) to the variance in total liability to definite schizophrenia.

By analyzing the rates of both definite and probable schizophrenia in the relatives of definite schizophrenics, we were able to conclude that the two forms of diagnosis were not manifestations of the same liability distribution. The two forms of diagnosis are strongly related, suggesting that a substantial portion of the liability underlying each diagnosis is shared in common. In addition to this common liability would be specific factors contributing toward the liability of each diagnosis. In analyzing the transmission of liability, we found some differences in the parameter estimates and results of hypotheses tests between the two forms of diagnosis. The major differences related to cultural transmission and marital resemblance, which are more marked for definite-plus-probable. It would appear that, should these differences be statistically reliable, those factors that differentiate the two forms of diagnosis are those environmental effects that spouses and other family members share in common [41].

For both forms of diagnosis, we found it necessary to posit some form of special twin resemblance. Although this might reflect greater twin environmental similarity, it could also reflect other factors that would increase twin concordance rates relative to the rates for other types of relatives. One such factor is genetic dominance, which would increase concordance rates in twins and siblings as compared to all other types of relatives. To test whether dominance was operating, we reparameterized the model by deleting t^2 and added a dominance term d^2, which contributes d^2 to the monozygotic (MZ) correlation and $d^2/4$ to both the dizygotic (DZ) and full-sib correlations. Under this new model, we found that for both forms of diagnosis the estimates of the dominance variance were near zero and nonsignificant and, for definite diagnosis, resulted in a poor fit of the general model to the data. It would appear, therefore, that the special twin resemblance is not a result of genetic dominance, although epistatic effects remain possible.

In the present analyses, we attempted to use as much information as possible by introducing a scaling factor for between-study heterogeneity. Alternatively, we could have approached the problem by performing analyses only upon ho-

mogeneous subsets of the total data set. In table 5, we present the estimates of genetic and cultural heritability, and the residual variance common to twins that we obtained when certain portions of the data were deleted. The final two entries are relevant to the present discussion. The first gives the estimates that result if we delete all information on classes of relatives for which there is significant between-study variability. The estimated heritabilities are in close agreement with those determined from the total data set. The final entry gives the estimates based upon a single comprehensive investigation (Bleuler [1]; definite diagnosis only). Bleuler reported rates for spouses, children, siblings, half-siblings, nieces and nephews, and grandchildren. These relationships allow identification of all parameters of the general model except t^2, which requires observations upon twins. The point estimate for cultural heritability is larger and genetic heritability smaller than for the pooled data set, although the associated standard errors are sufficiently large so as to allow the alternative estimates to be consistent with the same underlying parameter values. If we compare the results for Bleuler with the results for the full data set deleting twin data, we find close agreement in the

TABLE 5

EFFECTS OF DELETING FAMILY DATA UPON THE ESTIMATES OF GENETIC AND CULTURAL HERITABILITIES, AND RESIDUAL VARIANCE COMMON TO TWINS

RELATIONSHIP DELETED	DEFINITE			DEFINITE-PLUS-PROBABLE		
	h^2	c^2	t^2	h^2	c^2	t^2
None	.668 ± .052	.191 ± .038	.141 ± .034	.628 ± .073	.286 ± .106	.086 ± .025
Marital	.679 ± .067	.187 ± .048	.134 ± .040	.614 ± .067	.284 ± .075	.101 ± .039
Children	.696 ± .066	.146 ± .047	.145 ± .048	.631 ± .083	.167 ± .124	.099 ± .041
Siblings	.707 ± .122	.191 ± .313	.102 ± .037	.647 ± .286	.286 ± .105	.054 ± .043
MZ twins	.669 ± .036	.195 ± .026	.136 ± .025	.546 ± .200	.377 ± .233	.077 ± .039
DZ twins	.615 ± .053	.208 ± .030	.176 ± .033	.601 ± .149	.295 ± .179	.104 ± .083
Half-siblings	.671 ± .117	.201 ± .304	.128 ± .063	.637 ± .080	.281 ± .140	.082 ± .041
First cousins	.679 ± .100	.187 ± .389	.134 ± .058	.651 ± .083	.275 ± .144	.074 ± .044
Nieces-nephews	.678 ± .126	.191 ± .431	.131 ± .068	.629 ± .072	.277 ± .133	.081 ± .038
Grandchildren	.680 ± .094	.184 ± .405	.136 ± .054	.604 ± .082	.304 ± .134	.092 ± .042
All second degree	.668 ± .183	.207 ± .423	.125 ± .092	.643 ± .079	.277 ± .154	.080 ± .041
All second and third degree	.668 ± .156	.208 ± .353	.124 ± .082	.677 ± .028	.250 ± .023	.073 ± .018
All twin	.596 ± .188	.404 ± .308	*	.417 ± .201	.530 ± .227	*
All heterogeneous (cf. table 1)	.695 ± .066	.217 ± .067	.088 ± .038	.653 ± .075	.249 ± .093	.038 ± .043
All studies *other than* Bleuler [1]	.604 ± .339	.396 ± .534	*	†	†	*

* t^2 not estimable.

† Probable diagnosis was not used for children in Bleuler [1] study.

parameter estimates. As can be seen in table 5, it is only the deletion of the twin data that has marked influence upon the heritability estimates, this effect being greater for definite-plus-probable as compared with definite. The effect is to keep the genetic heritability higher and the cultural heritability lower than it would otherwise be had twin data not been included.

With a few exceptions, our results for definite diagnosis are in close agreement with those of Rao et al. [8] based on a similar data set for definite diagnosis. Although Rao et al. found significant cultural transmission while we did not, their estimate of .203 ± .120 for c^2 is in close agreement with ours, and their finding of significance could be a result of not correcting for between-study heterogeneity (a factor that these researchers noted could lead to spurious conclusions). The second major difference involves marital resemblance. Rao et al. found significant and substantial marital resemblance on familial environmental effects (estimated $u = .790$), while we found a nonsignificant and moderate amount of marital resemblance ($u = .246 \pm .328$). This latter discrepancy is, most probably, a result of our including two additional studies of the spouses of schizophrenics for which the reported risks were quite low.

The multifactorial threshold model for the transmission of schizophrenia is consistent with existing family data on schizophrenics. As important as fitting existing data is the framework provided by the multifactorial models for future research efforts. Recent advances in the study of environmental indices and the transmission of correlated phenotypes [42, 43] will hopefully be extended to multifactorial models for qualitative phenotypes. If so, such models could provide a powerful means to investigate putative environmental and biological liabilities for schizophrenia.

ACKNOWLEDGMENTS

We gratefully acknowledge the many valuable comments and criticisms of C. R. Cloninger and R. C. Elston.

APPENDIX
EFFECT OF BETWEEN-STUDY HETEROGENEITY UPON THE VARIANCE OF A POOLED PROPORTION

NOTATION

Assume that for a given type of relationship, we observed sample proportions, $p_1, p_2, \ldots, p_k$, from each of k independent studies. Let s_i be the arc sin transform of the ith sample proportion, then the distribution of s_i is, approximately [44], $s_i \sim N(\theta_i, 1/4Ni)$, where $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2 and N_i is the sample size observed in the ith study.

If $\theta = \theta_1 = \theta_2 = \ldots = \theta_k$, then there is no between-study heterogeneity. In general, assume $\theta_i \sim N(\theta, \sigma^2)$, where θ is the common parameter and σ^2 is the variance due to heterogeneity. If s equals the sample-size-weighted pooled estimate of θ (i.e.,

$$s = \sum_{i=1}^{k} N_i s_i / N$$

where

$$N = \sum_{i=1}^{k} N_i$$

is the total sample size), then what we require is an estimate of the variance of s under the general situation when there is between-study heterogeneity.

DISTRIBUTION AND EXPECTATION OF THE HETEROGENEITY χ^2

Under the formulation given above it is easy to show that the marginal distribution of s_i is $N(\theta, \sigma^2 + 1/4N_i)$, so that the pooled estimate, s, has distribution $N(\theta, 1/4N\ (1 + 4\sigma^2\ \Sigma N_i^2/N)$, the term $(1 + 4\sigma^2\ \Sigma N_i^2/N)$ reflecting the inflation of variance due to heterogeneity. The heterogeneity χ^2 is defined as

$$\chi^2 = \sum_{i=1}^{k} 4N_i(s_i - s)^2 ,$$

which can be expressed as

$$\chi^2 = \sum_{i=1}^{k} 4N_i(s_i - \theta)^2 - 4N(s - \theta)^2$$

$$= \sum_{i=1}^{k} 4N_i(\sigma^2 + 1/4N_i)\chi^2_1 - (1 + 4\sigma^2\Sigma N_i^2/N)\chi^2_1$$

$$= \chi^2_{k-1} + (k - 1)4\sigma^2 \frac{\Sigma N_i^2}{N} - 4\sigma^2 k \frac{\text{var}(N_i)}{\bar{N}} ,$$

where $\text{var}(N_i)$ is the variance of the observed sample sizes and $\bar{N}$ is the average sample size. Dividing both sides by the degrees of freedom and taking expectations gives

$$E(\chi^2/k - 1) = \left\{1 + 4\sigma^2 \frac{\Sigma N_i^2}{N}\right\} - 4\sigma^2\left(\frac{k}{k - 1}\right)\left(\frac{\text{var}(N_i)}{\bar{N}}\right) .$$

The braced term is precisely the inflation factor for the variance in the pooled estimate s. Consequently, using a total scaled sample size of $N \div (\chi^2/k - 1)$ will give a better approximation to the variance of s, especially when the variance in sample sizes is small relative to the mean (i.e., when the final term in the above equation is small).

REFERENCES

1. BLEULER M: *The Schizophrenia Disorders: Long-term Patient and Family Studies*. New Haven, Conn., Yale Univ. Press, 1978
2. GOTTESMAN II, SHIELDS J: *Schizophrenia: The Epigenetic Puzzle*. New York, Cambridge Univ. Press, 1982
3. DEBRAY Q, CAILLARD V, STEWART J: Schizophrenia: a study of genetic models and some of their implications. *Neuropsychobiology* 4:217–219, 1978

4. DEBRAY Q, CAILLARD V, STEWART J: Schizophrenia: a study of genetic models. *Hum Hered* 29:27–36, 1979
5. ELSTON RC: Discussion. Methodologies in human behavior genetics. *Soc Biol* 20:276–279, 1973
6. MATTHYSSE SW, KIDD KK: Estimating the genetic contribution to schizophrenia. *Am J Psychiatry* 133:185–191, 1976
7. O'ROURKE DH, GOTTESMAN II, SUAREZ BK, RICE J, REICH T: Refutation of the general single-locus model for the etiology of schizophrenia. *Am J Hum Genet* 34:630–649, 1982
8. RAO DC, MORTON NE, GOTTESMAN II, LEW R: Path analysis of qualitative data on pairs of relatives: application to schizophrenia. *Hum Hered* 31:325–333, 1981
9. ROSANOFF AS, ORR FI: A study of heredity in insanity in the light of the Mendelian theory. *Am J Insanity* 68:221–261, 1911
10. SLATER E, COWIE V: *The Genetics of Mental Disorders*. London, Oxford Univ. Press, 1971
11. ZERBIN-RÜDIN E: Endogene psychosen, in *Humangenetik*, vol 2, edited by BECKER PE, Stuttgart, Thieme, 1977, pp 446–577
12. ZERBIN-RÜDIN E: Psychiatrische genetik, in *Psychiatrie der Gegenwart*, vol 1, 2nd ed, edited by KISKER KP, MEYER JE, MÜLLER C, STRÖMGREN E, Berlin, Springer-Verlag, 1980, pp 545–618
13. KIDD KK, CAVALLI-SFORZA LL: An analysis of the genetics of schizophrenia. *Soc Biol* 20:254–265, 1973
14. TSUANG MT, BUCHER KD, FLEMING JA: Testing the monogenic theory of schizophrenia: an application of segregation analysis to blind family study data. *Br J Psychiatry* 140:595–599, 1982
15. FALCONER DS: The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet* 29:51–76, 1965
16. GOTTESMAN II, SHIELDS J: A polygenic theory of schizophrenia. *Proc Natl Acad Sci USA* 58:199–205, 1967
17. MORTON NE, MACLEAN CJ: Analysis of family resemblance. III. Complex segregation analysis of quantitative traits. *Am J Hum Genet* 26:489–503, 1974
18. CARTER CL, CHUNG CS: Segregation analysis of schizophrenia under a mixed genetic model. *Hum Hered* 30:350–356, 1980
19. SHIELDS J, GOTTESMAN II: Cross-national diagnosis of schizophrenia in twins. *Arch Gen Psychiatry* 27:725–730, 1972
20. ENDICOTT J, NEE J, FLEISS J, COHEN J, WILLIAMS JBW, SIMON R: Diagnostic criteria for schizophrenia. *Arch Gen Psychiatry* 39:887–889, 1982
21. RICE J, NICHOLS P, GOTTESMAN II: Assessment of sex differences for qualitative multifactorial traits using path analysis. *Psychiatr Res* 4:301–312, 1981
22. KETY SS, ROSENTHAL D, WENDER PH, SCHULSINGER F: Studies based on a total sample of adopted individuals and their relatives: why they were necessary, what they demonstrated and failed to demonstrate. *Schizophrenia Bull* 2:413–428, 1967
23. ESSEN-MÖLLER E: The calculation of morbid risk in parents of index cases as applied to a family sample of schizophrenics. *Acta Genet Stat Med (Basel)* 5:334–342, 1955
24. BLEULER E (translated by ZINKIN J): *Dementia Praecox or the Group of Schizophrenias*. New York, International Universities Press, 1950 (originally published 1911)
25. GOTTESMAN II, SHIELDS J: *Schizophrenia and Genetics*. New York, Academic Press, 1972
26. KRINGLEN E: *Heredity and Environment in the Functional Psychoses*. London, Heinemann, 1967
27. SLATER E (with the assistance of SHIELDS J): *Psychotic and Neurotic Illnesses in Twins*. Medical Research Council Special Report Series No. 278. London, Her Majesty's Stationery Office, 1953
28. RAO DC, MORTON NE: IQ as a paradigm in genetic epidemiology, in *Genetic Epidemiology*, edited by MORTON NE, CHUNG CS, New York, Academic Press, 1978, pp 145–182

29. RAO DC, MORTON NE, LALOUEL JM, LEW R: Path analysis under generalized assortative mating. II. American I.Q. *Genet Res (Camb)* 39:187–198, 1982
30. REICH T, JAMES JW, MORRIS CA: The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. *Ann Hum Genet* 36:153–184, 1972
31. ESSEN-MÖLLER E: Individual traits and morbidity in a Swedish rural population. *Acta Psychiatr Neurol Scand Suppl* 100, 1956
32. FREMMING K: *The Expectation of Mental Infirmity*. Occasional Papers on Eugenics No. 7. London, Cassell, 1951
33. HELGASON T: Epidemiology of mental disorders in Iceland. *Acta Psychiatr Scand [Suppl]* 173, 1964
34. SMITH C: Statistical resolution of genetic heterogeneity in familial disease. *Ann Hum Genet* 39:281–291, 1976
35. MORTON NE, RAO DC, LALOUEL JM: *Methods in Genetic Epidemiology*. New York, S. Karger, 1983
36. RAO DC, MORTON NE, CLONINGER CR: Path analysis under generalized assortative mating. I. Theory. *Genet Res* 33:175–188, 1979
37. KARLSSON JL: Family transmission of schizophrenia: a review and synthesis. *Br J Psychiatry* 140:600–606, 1982
38. PAYKEL ES: Contribution of life events to causation of psychiatric illness. *Psychol Med* 8:245–253, 1978
39. SCHOFIELD W, BALIAN L: A comparative study of the personal histories of schizophrenic and nonpsychiatric patients. *J Abnorm Soc Psychol* 59:216–225, 1959
40. RABKIN JG: Stressful life events and schizophrenia: a review of the research literature. *Psychol Bull* 87:408–425, 1980
41. VAUGHN C, LEFF JP: The influence of family and social factors on the course of psychiatric illness. *Br J Psychiatry* 129:125–137, 1976
42. RAO DC, MCGUE M, WETTE R, GLUECK CJ: Path analysis in genetic epidemiology, in *Methods in Human Population Genetics*, edited by CHAKRAVARTI A, Stroudsburg, Pa., Hutchinson Ross, 1983
43. MCGUE M, RAO DC, REICH T, LASKARZEWSKI P, GLUECK CJ, RUSSELL JM: The Cincinnati Lipid Research Clinic Family Study: bivariate path analyses of lipoprotein concentrations. In preparation
44. BISHOP YMM, FIENBERG SE, HOLLAND PW: *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Mass., MIT Press, 1975